

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-352521

(43)Date of publication of application : 21.12.2001

(51)Int.Cl.

H04N 5/928
G10L 19/00
G11B 20/10
G11B 20/12
G11B 27/034
H04N 5/765
H04N 5/781
H04N 5/91
H04N 5/92
H04N 7/24

(21)Application number : 2000-177132

(71)Applicant : HITACHI LTD

(22)Date of filing : 08.06.2000

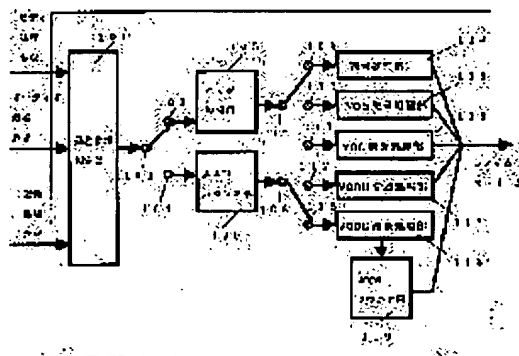
(72)Inventor : IMAI TSUTOMU
YOSHIDA SUSUMU
TARUMI HIROYUKI

(54) INFORMATION RECORDER

(57)Abstract:

PROBLEM TO BE SOLVED: To solve the problem of a conventional information recorder that has difficulty in seamless reproduction, partial erasure and edit because a position of an audio pack of an audio stream with respect to a position of a video pack of a video stream is at random in the case of an MPEG system stream.

SOLUTION: This invention realizes an image audio compressor that generates a system stream facilitating partial erasure, edit and seamless reproduction at a high-speed by always multiplexing video data by 1 GOP in 1 VOB and audio data in a period equivalent to that of the video data.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the

THIS PAGE BLANK (USPTO)

examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

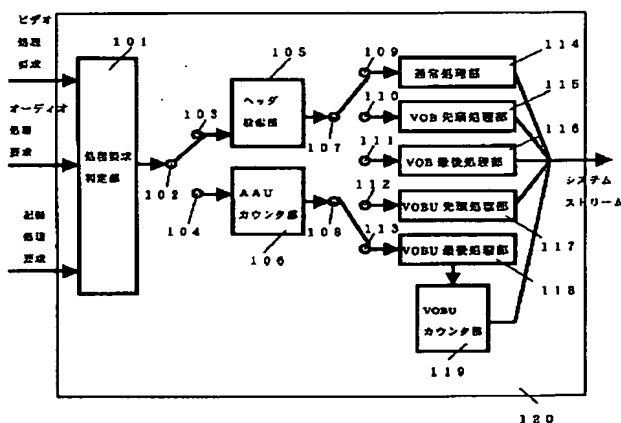
[Number of appeal against examiner's decision of
rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

THIS PAGE BLANK (USPTO)



【特許請求の範囲】

【請求項 1】 レンズにより集光された光を結像し電気信号に変換する撮像素子と、該電気信号を映像信号に変換するカメラ信号処理部と、該映像信号を圧縮しビデオビットストリームを生成するビデオ処理部と、集音手段で集められた音声をデジタル信号に変換する A/D 変換部と、デジタル信号に変換されたオーディオデータを圧縮してオーディオビットストリームを生成するオーディオ処理部と、上記ビデオ処理部で生成されたビデオストリームと上記オーディオ処理部で生成されたオーディオストリームとを同期を取りシステムストリームを生成する多重化部と、該システムストリームを情報記録媒体に記録する記録手段とを備えていることを特徴とする画像音声圧縮装置。

【請求項 2】 請求項 1 に記載されている多重化部は、処理要求判定部、スイッチ、ヘッダ検索部、AAU カウンタ、通常処理部、VOB 先頭処理部、VOB 最後処理部、VOBU 先頭処理部、VOBU 最後処理部、VOBU カウンタ部とから構成することを特徴とする画像音声圧縮装置。

【請求項 3】 請求項 2 に記載されている多重化部が生成するシステムストリームは 1 VOB 内に 1 GOP のビデオストリームを多重化し、多重化したビデオストリームと同時間のオーディオストリームを多重化することを特徴とする画像音声圧縮装置。

【請求項 4】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、VOBU 内でのビデオパックとオーディオパックの間隔はビデオのビットレートとオーディオのビットレートから算出し、常に等間隔で多重化することを特徴とする画像音声圧縮装置。

【請求項 5】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、オーディオストリームは 21 AAU を 5 VOB 20 AAU を 1 VOB の 6 VOB 周期で多重化することを特徴とする画像音声圧縮装置。

【請求項 6】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、オーディオストリームが MPEG Audio 方式で圧縮されている場合には 48 VOB 毎に 1 AAU 増やして多重化することを特徴とする画像音声圧縮装置。

【請求項 7】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、オーディオストリームが MPEG Audio 方式で圧縮されている場合には 48000 VOB 毎に 1 AAU 増やして多重化することを特徴とする画像音声圧縮装置。

【請求項 8】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、オーディオストリームが AC-3 方式で圧縮されている場合には 64 VOB 毎に 1 AAU 増やして多重化することを特徴とする画像音声圧縮装置。

【請求項 9】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、オーディオストリームが AC-3 方式で圧縮されている場合には 64000 VOB 毎に 1 AAU 増やして多重化することを特徴とする画像音声圧縮装置。

【請求項 10】 請求項 2 に記載されている多重化部が生成するシステムストリームにおいて、VOBU の先頭は必ずビデオパックを配置し、次にオーディオパックを配置することを特徴とする画像音声圧縮装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、符号化された動画像と音声とを記録媒体に記録する装置に関わり、特に動画像と音声とを同期を取って記録媒体に記録する画像音声圧縮装置に関するものである。

【0002】

【従来の技術】 MPEG 方式で圧縮された動画像と音声とを多重化する方法として ISO/IEC 13818-1 MPEG-2 SYSTEMS がある。以下、これを公知例 1 とする。公知例 1 では、タイムスタンプを使い、動画像と音声との同期をとる。

【0003】 動画像と音声とを同期を取って記録媒体、例えば DVD-RAM に記憶する方法として特開平 11-155131 がある。以下、これを公知例 2 とする。公知例 2 のシステムストリームの構成を図 3 を用いて説明する。

【0004】 301 から 303 は VOB (Video Object) と呼ぶ。ここで、301 の VOB #1 を例にとって説明する。VOB は記録開始から記録終了までの一連のシステムストリームである。301 の VOB は 304 から 306 の VOB U (Video Object Unit) から構成されている。VOBU とは、その再生時間が約 0.4 秒～1.0 秒となるピクチャデータからなる少なくとも 1 つ以上の GOP (Group of Pictures) と、このピクチャデータと共に多重化されるオーディオデータを含む単位である。304 の VOB U は、307 から 310 の V_PCK (Video Pack) と A_PCK とから構成されている。V_PCK (Video Pack)、A_PCK とともに 2048 バイト固定長である。

【0005】 公知例 2 の pp16 に記載されているように MPEG 規格では、バッファ内にデータを蓄積できる時間の上限が規定されており、全てのデータはバッファに入力されてから 1 秒以内にバッファから取り出されなければならないという制約の 1 秒ルールがある。(MPEG 規格の再生バッファモデルの詳細は発明の実施の形態を参照)

【0006】

【発明が解決しようとする課題】 公知例 2 には、VOB を n 個の VOB U にすることやその利点について記載されているが、映像信号と音声信号の多重化については記載されていない。

【0007】 また、上記に記載した 1 秒ルールのため

に、システムストリームでのビデオデータとオーディオデータは極端に離れて多重化されるということはない。MPEG-2で圧縮されるシステムストリームのビットレートでは、通常6Mbps程度で圧縮される。高画質で圧縮される場合には10Mbpsを越えて圧縮される場合もある。そのため、ビデオストリームとオーディオストリームとが1秒離れて多重化されていた場合にはそのストリームを再生するのに多大なバッファを必要とする。

【0008】また、仮に公知例2のように1VOBU毎に時間を区切った上で、単純に映像信号や音声信号を多重化した場合、1VOBU内のビデオストリームのビデオパックに対するオーディオストリームのオーディオパック位置はまちまちとなってしまうため、シームレス再生、部分削除、編集作業等を簡単に行うことは困難である。

【0009】本発明では、VOBのような一続きのストリームをVOBUのような複数のストリームに分割した上でビデオデータとオーディオデータを多重化できるようにすることや、編集を容易にできるようにすることをその目的とする。

【0010】

【課題を解決するための手段】本発明では、1VOBU内で、1GOP分のビデオデータに、1AAU分のオーディオデータを、1GOP分の時間とほぼ同じ時間に相当する数だけ多重化するようにする。このようにすることで、容易かつ高速に部分削除、編集およびシームレス再生可能なシステムストリームを生成する画像音声圧縮装置を実現するものである。

【0011】

【発明の実施の形態】以下、本発明の実施例を図面を用いて説明する。まず、本発明の実施例の動作について説明する。図2は、本発明の一実施例のシステムブロック図である。レンズ201は、光を集光する。撮像素子202は、レンズ201で集光された光を結像し、撮像信号（電気信号）に変換する撮像素子である。カメラ処理部203は、撮像素子202より得られる撮像信号をテレビジョン等で写す為の映像信号に変換する。ビデオ処理部204は、カメラ信号処理部203の映像信号を圧縮しビデオビットストリームを生成する。マイク205は、音を集音する。A/D変換部206は、マイク205で集められた音声デジタル信号に変換する。オーディオ処理部207は、A/D変換部206でデジタルに変換されたオーディオデータを圧縮してオーディオビットストリームを生成する。多重化部208は、ビデオ処理部204で生成されたビデオストリームとオーディオ処理部207で生成されたオーディオストリームとを同期を取りシステムストリームを生成する。マイコン209は、オーディオ処理部207と多重化部208とで構成される。記録媒体210には、映像と音声のシステム

ストリームが記録される。ボタン211は記録開始と記録停止のボタンである。これらの構成要素により、ビデオデータとオーディオデータを取り込み、圧縮し、システムストリームを生成して、記録媒体210としてのDVD-RAMに記録するDVDカメラ211を構成する。

【0012】映像はレンズ201で集光され、撮像素子202により撮像信号に変換され、カメラ信号処理部203で映像信号に変換され、ビデオ処理部204でMPEG-2に準拠したビデオストリームに圧縮される。一方音声は、マイク205で集音されA/D変換部206でデジタルに変換されオーディオ処理部207でMPEG AudioもしくはAC-3に圧縮される。多重化部208でビデオ処理部204が生成したビデオストリームとオーディオ処理部207が生成したオーディオストリームとを同期を取りMPEGシステムストリームに準拠した形式でDVD-RAM等の記録媒体210に記録する。

【0013】次に、図2の多重化部208が生成するシステムストリームについて図1、図3、図10、図11、図12及び図17を用いて説明する。本発明のVOB、VOBUの構成は上記公知例2と同じである。

【0014】まず、多重化部208の構成を図1を用いて説明する。101は処理要求判定部、102から104はスイッチ、105はヘッダ検索部、106はAAUカウンタ、107から113はスイッチ、14は通常処理部、115はVOB先頭処理部、116はVOB最後処理部、117はVOBU先頭処理部、118はVOBU最後処理部、119はVOBUカウンタ部、120は多重化処理部である。

【0015】次に、多重化部208の処理手順を図17のフローチャートを用いて説明する。ビデオ処理部204またはオーディオ処理部207では、圧縮されたビデオストリームまたは圧縮されたオーディオストリームが一定量を越えると多重化部208に対し処理要求を発行する（ステップS101）。

【0016】次に、処理要求判定部101は、ビデオ処理部204からの処理要求かオーディオ処理部207からの処理要求かを判定し（ステップS102）、ビデオ処理要求の場合には、ヘッダ検索部105にヘッダ検索を実行させる（ステップS103）。一方、オーディオ処理要求の場合には、AAUカウンタ部104にAAUのカウンタを実行させる（ステップS104）。

【0017】なお、処理要求判定部101は、ボタン211が押された場合には記録処理要求が発行されたと判定し、記録処理を実行していない状態で記録処理要求が発行された場合には記録開始処理要求と判定し、記録中に記録処理要求が発行された場合には記録停止処理要求と判定する。

【0018】まず、ビデオ処理要求の場合を説明する。

ヘッダ検索部105ではピクチャヘッダ、シーケンスヘッダまたはGOPヘッダを検索する。そして、ステップS105にて、ヘッダ検索部105が、VOBの先頭か否かをチェックする。処理要求判定部101が記録開始処理要求と判定した場合には、VOBの先頭処理部115にてVOB先頭処理を実行する(ステップS106)。VOBの先頭でない場合には、ステップS107にて、ヘッダ検索部105が、VOBの最後か否かをチェックする。ここでは処理要求判定102が記録停止処理要求と判定した場合に、VOB最後処理部111がVOB最後処理を実行する(ステップS108)。VOBの最後でない場合にはステップS109にて、ヘッダ検索部105が、VOBUの先頭か否かをチェックする(ステップS109)。

【0019】記録動作中にヘッダ検索部105でシーケンスヘッダまたはGOPヘッダを見つけた場合、その検索したヘッダ位置が読み出そうとしているビデオストリームの先頭にある場合には、VOBU先頭処理部112にVOBU先頭処理を実行させる(ステップS110)。一方、先頭でない場合には、ステップS111に
20 移行し、ヘッダ検索部105が、VOBUの最後か否かをチェックする。そして記録動作中にヘッダ検索部105でシーケンスヘッダまたはGOPヘッダを見つけた場合であって、見つけたヘッダ位置が読み出そうとしているビデオストリームの先頭でない場合には、VOBU最後処理部118が、VOBU最後処理を実行する(ステップS112)。さらに、ビデオ処理要求のVOBUカウンタ119でVOBUカウンタ処理を行い、カウンタをインクリメントする(ステップS113)。上記のどの条件にも当てはまらない場合には、通常処理114を
30 実行して終了となる(ステップS115)。

【0020】次にオーディオ処理要求の場合について説明する。処理要求がオーディオ処理要求であると処理要求判定部101が判断した場合には、AAUカウンタ部106が、AAUカウンタ処理を実行する(ステップS104)。AAUカウンタ部106は、VOBU毎にクリアするカウンタ(VOBU内AAUカウンタ)とVOBの先頭からカウントして記録停止でクリアするカウンタ(VOB内AAUカウンタ)の2つのカウンタで構成される。

【0021】ステップS105にて、AAUカウンタ部106がVOB内AAUカウンタをチェックして、カウンタ値が零の場合には、VOB先頭処理部115がVOB先頭処理を実行する(ステップS106)。VOBの先頭ではない場合には、ステップS107に移行し、AAUカウンタ部106がVOBの最後か否かをチェックする。ここでは処理要求判定102が記録停止処理要求と判定した場合にVOB最後処理108を実行する。

【0022】ステップS109では、VOB内AAUカウンタが零でない場合であってVOBU内AAUカウン

タが零の場合に、VOBU先頭処理部112がVOBU先頭処理を実行する(ステップS110)。VOB内AAUカウンタとVOBU内AAUカウンタがともに零でない場合には、ステップS111に移行し、VOBU内AAUカウンタが21をカウントした場合には、VOBU最後処理部118が、VOBU最後処理を実行する(ステップS112)。さらに、オーディオ処理要求のVOBUカウンタ113でカウンタをインクリメントする(ステップS113)。上記のどの条件にも当てはまらない場合には、通常処理114を実行して終了となる(ステップS115)。

【0023】次に、通常処理部114、VOB先頭処理部115、VOB最後処理部116、VOBU先頭処理部117、VOBU最後処理部118における処理内容について説明する。

【0024】まず、通常処理部114が生成するシステムストリームについて図13を用いて説明する。一実施例として、ビデオストリームはNTSC方式、MPEG-2としビットレートを6Mbpsとする。オーディオストリームはMPEG Audioとしビットレートを
384kbps、サンプリング周波数48kHz、圧縮モードはLayer 2とする。

【0025】1VOBUの大きさは0.4秒から1.0秒であることが好ましい。そこで、本発明では、1VOBUの長さは0.5秒とした。ビデオストリームは、1VOBU=1GOP=15フレームとして1VOBUに多重化するためである。シームレス、編集、再生、特殊再生をより簡単に行うため、本発明では1VOBU内に
40 同じ時間のビデオストリームとオーディオストリーム多重化する。オーディオストリームは、1VOBU内ではAAUで完結しなければいけない。上記仕様で圧縮したオーディオストリームの1AAU(Audio Access Unit)は0.024秒である。しかし、AAUの倍数

(0.024秒の倍数)でしか1VOBUに多重化することができないため、オーディオストリームは20AAU($0.024 \times 20 = 0.48$ 秒)または21AAU($0.024 \times 21 = 0.504$ 秒)のどちらかを選択しなければならない。そこで、21AAUを5VOBUと20AAUを1VOBUとして6VOBU周期の構成とした。つまり、 $(0.024 \times 21 \times 1) + (0.024 \times 20 \times 5) = 3.00$ 秒分のオーディオストリームを多重化する。

【0026】ここで、新たな問題が生じる。NTSC方式ではビデオフレームレートが29.97Hzであるため1VOBU=0.5秒とはならない。1VOBUでは $1/29.97 \times 15 = 0.500500 \dots$ と端数が出てしまうため長時間記録した場合にはどんどん端数が積み重なってビデオストリームとオーディオストリームとが離れてしまう。

【0027】ここで、どれくらいずれるのかを算出す

る。

6VOBUでのずれは、 $1/29.97 \times 15 \times 6 = 3.00300300 \dots$

12VOBUでのずれは、 $1/29.97 \times 15 \times 12 = 6.00600600 \dots$

24VOBUでのずれは、 $1/29.97 \times 15 \times 24 = 12.0120120 \dots$

48VOBUでのずれは、 $1/29.97 \times 15 \times 48 = 24.0240240 \dots$

となり48VOBU(24秒分)でオーディオ1AAU 10 分以上となる。

【0028】そのため、48VOBU毎に1AAUを増やしてビデオストリームとオーディオストリームとのずれを吸収する。しかし、48VOBU毎に1AAU増やしてもまだ端数が生じているため、更に48000VOBU(24000秒=400分=約6.7時間)に一回1AAUを増やすことにより長時間記録の場合にも、ビデオデータとオーディオデータの端数を吸収するためビデオとオーディオが離れることはない。

【0029】次に、オーディオストリームがAC-3の 20 場合について説明する。AC-3の場合でサンプリング周波数が48kHzの場合には1AAUは0.032秒である。そのため、上記と同様の演算を行うと、

6VOBUでのずれは、 $1/29.97 \times 15 \times 6 = 3.00300300 \dots$

16VOBUでのずれは、 $1/29.97 \times 15 \times 16 = 8.00800800 \dots$

32VOBUでのずれは、 $1/29.97 \times 15 \times 32 = 8.01601601 \dots$

64VOBUでのずれは、 $1/29.97 \times 15 \times 16 = 8.03203203 \dots$ 30

となり64VOBU(32秒分)でオーディオ1AAU 分以上となる。

【0030】そのため、64VOBU毎に1AAUを増やしてビデオストリームとオーディオストリームとのずれを吸収する。しかし、64VOBU毎に1AAU増やしてもまだ端数が生じているため、更に64000VOBU(32000秒=約533分=約8.9時間)に一回1AAUを増やすことにより長時間記録の場合にもビデオデータとオーディオデータの端数を吸収するためビデオとオーディオが離れることはない。

【0031】ここで、オーディオの圧縮方式がMPEG Audioの場合についてさらに詳しく説明する。上記のように6VOBU毎の周期でビデオストリームは1VOBU=1GOP、オーディオストリームは5VOBU=21AAU、1VOBU=20AAUとなるように構成する。ビデオパックとオーディオパックとの間隔はVOBUの先頭付近と最後付近を除いて等間隔で配置する。まず、この等間隔で配置する間隔の算出の方法について説明する。

【0032】まず、オーディオが21AAUの場合ビデオとオーディオの間隔について図13を用いて説明する。

【0033】ビデオストリームのビットレートを6Mbps、オーディオストリームのビットレートを384kbpsとすると $6\text{Mbps}/384\text{kbps} = 15.625$ となる。したがって、ビデオパックを15個連続した後にオーディオパックを1つとなるような構成とする。図13において、V1301からV1302が15パック連続した後にA1303が多重化され続いてV1304からV1305が15パック連続した後にA1306が多重化される。

【0034】ビデオパックとオーディオパックの詳細を図4、図5、図6、図7、図8、図14を用いて説明する。

【0035】まず、ビデオパックについて図4、図5を用いて説明する。

【0036】図4にビデオパックとオーディオパックの基本構成を示した。401はPack_Start_code、402はSCR、403はProgram_Mux_rate、404はPack_stuffing_lengthとから構成されるPack_header(14バイト)とビデオパケットまたはオーディオパケットとから構成される。1パックは2048Bの固定長である。

【0037】図5にビデオパックの様子を示した。501の上図はVOBUの先頭でのビデオパックを示してあり、下図はそれ以外のビデオパックを示してある。501はPack_Header、502はSystem header、503のPack_headerと504のvideo dataでV_PKT(Video Packet)を構成している。VOBUの先頭以外では502のSystem headerがない構成である。つまり、505のPack_Header、506のPacketheader、507のVideo dataとから構成される。

【0038】次に、図6、図7、図8を用いてオーディオパックについて説明する。

【0039】図6は、オーディオの圧縮にMPEG Audio方式の場合でのオーディオパックの様子が示してある。601はPack_Header、602はPacket header、603はAudio dataである。

【0040】図7は、オーディオの圧縮にAC-3方式の場合のオーディオパックの様子が示してある。701はPack_Header、702はPacket header、703はSubstream_id、704はAudio_frame_information、705はAudio dataである。

【0041】図8は、オーディオを非圧縮するLinear PCMの場合のオーディオパックの様子が示してある。801はPack_Header、802はPacket header、803はSubstream_id、804はAudio_frame_information、805はAudio_data_information、806はAudio dataである。

【0042】ここでMPEG規格での再生バッファモデ 50

ルについて図9を用いて説明する。

【0043】図9において、901はSTC (System Time Clock)、902はDEMUX (分離部)、903はVideo Buffer、904はVideo decoder、905はAudio Buffer、906はAudio decoder、907は、再生バッファモデル、908はVideo出力、909はAudio出力である。Video BufferサイズはMPEG-2で232kB、MPEG-1で46kB、Audio Bufferサイズは4kBである。

【0044】ビデオとオーディオの同期にはシステムストリーム内のパックヘッダに含まれるSCR (System Clock Reference)、パケットヘッダに含まれるデコード時刻を示すDTS (Decoding Time Stamp)とデコードされたビデオストリームとオーディオストリームを何時出力するかを指示するPTS (Presentation Time Stamp)とから同期とる。

【0045】MPEG規格での再生バッファモデルは、SCRを用いて時刻基準となるSTCの値を多重化側で意図した値にセットされ、システムストリームのビットレートでDEMUX902に読み込まれる。SCRはMPEG-2では6バイト (実データ42ビット)、MPEG-1では5バイト (実データ33ビット) である。DEMUX902でビデオストリームおよびオーディオストリームに分離されVideo Buffer903とAudio buffer905に転送される。DTSとSTCが一致した時刻にビデオストリームはVideo decoder904、オーディオストリームはAudio decoder906に転送され伸長される。最後に、PTSとSTCが一致した時刻に伸長されたビデオデータはVideo出力908に出力され、伸長されたオーディオデータはAudio出力909に出力される。

【0046】本発明でのビデオパックとオーディオパック (MPEG Audio) のVOBUの先頭、PTS、DTSあり、PTSのみ、PTS、DTSなし、VOBUの最後において、Pack_headerのバイト数、System headerバイト数、packet headerのバイト数、stuffingのバイト数、padding packetのバイト数、およびVideo dataとAudio dataのバイト数を図14に示した。オーディオに関してはPTSとDTSが同じ値のためPTSのみをつける。また、パックには必ずAAUヘッダが含まれるため、図14にはVOBUの先頭とPTSのみを示してある。特長は、4バイト境界でのアクセスが可能となるようにStuffingでバイト数を調整している。これにより、メモリアクセスが4バイト境界になるため高速な多重化が実現できる。

【0047】次に、VOBU先頭処理部117とVOBU最後処理部118での多重化方法について図11を用いて説明する。

【0048】図11において、VOBUの切れ目はV1106とV1109との間とする。V1101からV1

106が前のVOBUであり、V1109以降が次のVOBUである。

【0049】まず、VOBU先頭処理部117での処理について説明する。

【0050】VOBUの先頭では、ビデオパックV1109を配置する。次にA1110を配置する。その後にV1111からV1112のビデオを配置する。上記具体例の場合には、ビデオパックとオーディオパックとの間隔は15パックとなるためV1111からV1112は15パックのビデオパックである。VOBUの先頭にビデオパックおよびオーディオパックを配置した理由は、VOBUの先頭に含まれるビデオパックとオーディオパックには伸長に必要な情報が含まれており、特殊再生やシームレス再生の場合にはその情報を基にビデオ処理部やオーディオ処理部を設定する必要がある。例えば、VOBUの先頭のビデオパックに含まれるシーケンスヘッダには画像の横の画素数、画像の縦のライン数、アスペクト比等であり、オーディオパックにはレイヤ、エラーチェック (CRC)、ビットレート、サンプリング周波数等である。そのため、素早く伸長情報を得るためにVOBUの先頭に配置した。また、特殊再生やシームレス再生の場合でVOBUのつなぎ目でビデオやオーディオの圧縮モードが変わる場合も考えられる。VOBUの先頭をビデオパックとし次をオーディオパックと固定することにより、伸長モード設定のためにVOBUの先頭付近での先頭ビデオパックと先頭オーディオパックを検索する時間も短縮することができる。

【0051】次に、VOBU最後処理部118での処理について説明する。公知例3では、ビデオストリームもオーディオストリームもVOBU内で完結しなければならない。そのため、VOBUの最後のA1102ではA_PCT1103とP_PCT1104の1パック2パケット構成とする。ビデオストリームに関しても同様にVOBUの最後のビデオパック1106ではV_PCT1107とP_PCT1108の1パック2パケット構成とする。

【0052】ここで、1VOBU内にビデオパックとオーディオパックがいくつ含まれるかについて算出する。1パックあたりの平均データ数を2020バイトとして算出すると

ビデオパック数： $6\text{Mbps} / 8 \times 0.5\text{秒} / 2020 = 185.64\dots$ となるためビデオパック数は186パック、オーディオパック数は1VOBUに21AAUまたは20AAU多重化されるため、ビットレート384kbpsの場合には1AAU=1152バイトであるので

21AAUの場合： $21 \times 1152 / 2020 = 11.97\dots$

20AAUの場合： $20 \times 1152 / 2020 = 11.40\dots$

となるため21AAUの場合も20AAUの場合でも1VOBUに多重化されるオーディオパック数は12パックである。

【0053】1VOBUでのビデオパックとオーディオパックの合計は186+12=198パックとなる。

【0054】オーディオパックは12パックなので198-(15+1)×12=6パック。VOBUの先頭はビデオパックなので1パック減らすと、VOBUの最後のビデオパックV1105からV1106のパック数は5パックとなる。

【0055】次に、VOB先頭処理部115が生成するシステムストリームについて図9、図10を用いて説明する。

【0056】VOBの先頭のオーディオパックのPTS(DTS)はビデオストリームの先頭のBピクチャと一致する。ビデオストリーム先頭のIピクチャのDTSをDTSi、ビデオストリーム先頭のBピクチャのDTSをDTSbとする。図9のDTSのクロックは90kHz、Video buffer903は232kBである。

【0057】ここで、Bピクチャの先頭のDTSbを算出する。ビデオフレームレートが29.97Hzの場合にはDTSb=DTSiとなるので

21AAUの場合：

$$\begin{aligned} \text{SCRの増分} &= 21 \times 2160 / 198 \times 27\text{MHz} / 90\text{kHz} \\ &= 68727 \end{aligned}$$

20AAUの場合：

$$\begin{aligned} \text{SCRの増分} &= 20 \times 2160 / 198 \times 27\text{MHz} / 90\text{kHz} \\ &= 65455 \end{aligned}$$

となる。

【0060】次に、VOBの先頭でのオーディオの出現位置と出現するまでのSCRvの増分を求める。ビデオのみのパックでのSCRvの増分は、

$$\begin{aligned} \text{SCRvの増分} &= 2048 \times 8 \times 27\text{M} / 6\text{M} \\ &= 73728 \end{aligned}$$

また、出現位置は、

$$6861300 / 73728 = 93.06$$

となるためVOBの先頭から94番目に最初のオーディオパックを配置する。

【0061】最初のVOBU内での残りビデオパックは186-93=93となる。ビデオとオーディオの間隔は15パックなので93/15=6.2であるので最初のVOBのVOBUにはオーディオパックを6パック配置する。6パックに含まれるAAUの数を算出すると1パック当たり2020バイトとして、

$$6 \times 2020 / 1152 = 10.52$$

となるため10AAU分のオーディオストリーム含める。

【0062】この構成は20AAUの後半と全く等しい構成となるため、SCRの増分は20AAUの場合と同じくとするればよい。

【0063】最後にVOB最後処理部116の生成方法

$$\begin{aligned} \text{DTSb} &= 232 \times 1024 \times 8 \times 90\text{k} / 6\text{M} + 3003 \\ &= 31511.16 \end{aligned}$$

したがって、

$$\begin{aligned} \text{DTSa} &= \text{DTSb} - (0.024 \times 90\text{k} \times 4) = 22871 \end{aligned}$$

SCRのクロックは27MHzなのでSCRに換算すると

$$22871 \times 27\text{M} / 90\text{k} = 6861300$$

ここで1パック当たりのSCRの増分を算出する。本実施例では演算量を削減するため、毎パックSCRを算出するのではなくVOBU毎に算出する。VOBUでのSCR増加量は一定である。1VOBUに含まれるオーディオAAUが21AAUと20AAUの場合があるためそれぞれについて算出する。

【0058】1VOBU当たりのSCRの増分は、SCRの増分=AAU数×1AAU当たりのPTSの増分×27MHz/90kHzである。

【0059】1AAU当たりのPTSの増分は、1AAU=0.024Sから、

$$0.024[\text{S}] \times 90\text{k}[\text{Hz}] = 2160$$

したがって1パック当たりの増分は、

を、図12を用いて説明する。

【0064】VOBの最後のVOBUも通常のVOBUの最後と同様にビデオストリームはGOPの切れ目、オーディオストリームはAAUの切れ目で完結しなければならない。

【0065】図12において1204はVOB最後のビデオパックを示しており1204はビデオパケットV_PCT1205とパディングパケットP_PCT1206とから構成される。1207から1208は連続したオーディオパックである。1208はオーディオパケットA_PCT1209とパディングパケット1210とから構成される。

【0066】本実施例では1VOBU内に含まれるビデオストリームとオーディオストリームとは時間的に常にオーディオが少々遅れて多重化されている。上記したVOBの先頭の分だけ遅れているので、本実施例では1207から1208のオーディオパック数は10パックとなる。

【0067】シームレス再生に関しても、VOBの先頭のVOBUではオーディオパックが10パック遅れているがVOBの最後のVOBUでは10パック多く多重化されているため、VOB毎のシームレス再生もスムーズに実現することができる。

【0068】上記では、ビデオのビットレートを固定し

た例を説明したが可変レートに関しても適応できることは明らかである。

【0069】次に、本発明で生成したシステムストリームを用いたシームレス再生と編集について説明する。

【0070】シームレス再生および編集の場合には、図9に示したVideo buffer 903とAudio buffer 905の2つのバッファに対してオーバーフロー、アンダーフローが起きないようにしなければならない。通常はVOBの切れ目やVOBUの切れ目でのビデオストリームとオーディオストリームはまちまちであるため、オーバーフローやアンダーフローを起こさずにシームレス再生を実現するには繋ぎ目でオーバーフローやアンダーフローを起こさないようにバッファに送り込まなければいけない。また、場合によっては編集のときには再多重化、再エンコードする必要が生じる。本発明で生成するシステムストリームは常にビデオストリームとオーディオストリームの位置関係が一定であるために容易に実行できる。

【0071】まず、部分削除について図15を用いて説明する。

【0072】図15において、(1)図は部分削除する前のVOBの様子であり、(2)図は部分削除した後のVOBの様子を示している。

【0073】(1)図において太線で囲んだ1503と1504の2つのVOBUを削除する場合について説明する。本発明で生成したシステムストリームではどの部分のVOBUでもビデオストリームに対するオーディオストリームの時間間隔が常に一定である。そのため、ビデオビットレート6Mbpsでオーディオビットレート384kbpsの場合には、上記したようにビデオストリームに対してオーディオストリームは10パケット分遅れている。(1)図において、削除されるVOBU1503にはVOBU1502のビデオストリームに対するオーディオストリームが10パケット含まれている。一方、VOBU1505には削除されるVOBU1504のビデオストリームに対するオーディオデータが同様に10パケット含まれている。したがって、VOBU1503に含まれているオーディオ10パケットをそのまま1505に移せば(2)図に示す編集後のシステムストリームが完成することになる。

【0074】次にVOBの編集について図11、図16を用いて説明する。

【0075】図16において(1)図は編集前の前半のVOB、(2)図は編集前の後半のVOB、(3)図は編集後のVOBを示している。(3)図において番号が(1)図および(2)図と同じ場合には同じパケットを示している。

【0076】上記でも説明したとおり、本発明で生成したシステムストリームではどの部分のVOBUでもビデオストリームとオーディオストリームとの時間間隔が常

に一定である。そのため、ビデオビットレート6Mbpsでオーディオビットレート384kbpsの場合には、上記したようにVOBの最後((1)図での1605から1606)はオーディオが10パケット連続する。ビデオストリームの繋ぎ目は(3)図での1604と1607である。この繋ぎ目をVOBUの切れ目とする。VOBUの切れ目以降では図11で説明したとおり先頭に1607を配置し、次にオーディオ1605を配置することにより実現する。

【0077】

【発明の効果】本発明では、1VOBU内にビデオデータを1GOP、オーディオデータをビデオデータに相当する時間を常に多重化することにより、容易に編集およびシームレス再生可能なビデオデータとオーディオデータとの間隔が一定となるシステムストリームを生成する画像音声圧縮装置を提供するものである。

【図面の簡単な説明】

【図1】多重化部の一実施例を示す図である。

【図2】画像音声圧縮装置の一実施例を示す図である。

【図3】システムストリームの構成を説明するための図である。

【図4】パックの構成を説明する図である。

【図5】ビデオパックの構成を説明する図である。

【図6】オーディオパックの構成を説明する図である。

【図7】オーディオパックの構成を説明する図である。

【図8】オーディオパックの構成を説明する図である。

【図9】MPEG規格での再生バッファモデルを説明する図である。

【図10】VOB先頭処理部の生成するシステムストリームを説明する図である。

【図11】VOBU先頭処理部とVOBU最後処理部の生成するシステムストリームを説明する図である。

【図12】VOB最後処理部の生成するシステムストリームを説明する図である。

【図13】通常処理部の生成するシステムストリームを説明する図である。

【図14】ビデオパックとオーディオパックの構成を説明する図である。

【図15】システムストリームの部分削除を説明する図である。

【図16】システムストリームの編集を説明する図である。

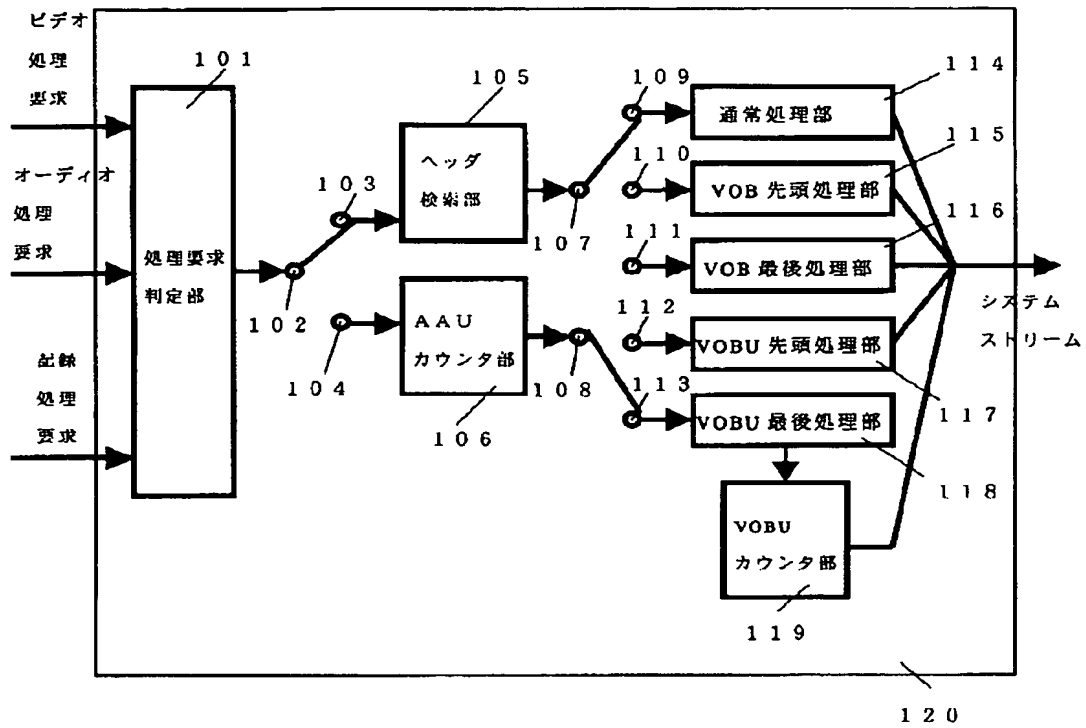
【図17】多重化部の処理手順を示す図である。

【符号の説明】

101…処理要求判定部、105…ヘッダ検索部、106…AAUカウンタ、114…通常処理部、115…VOB先頭処理部、116…VOB最後処理部、117…VOBU先頭処理部、118…VOBU最後処理部、119…VOBUカウンタ部、120…多重化部。

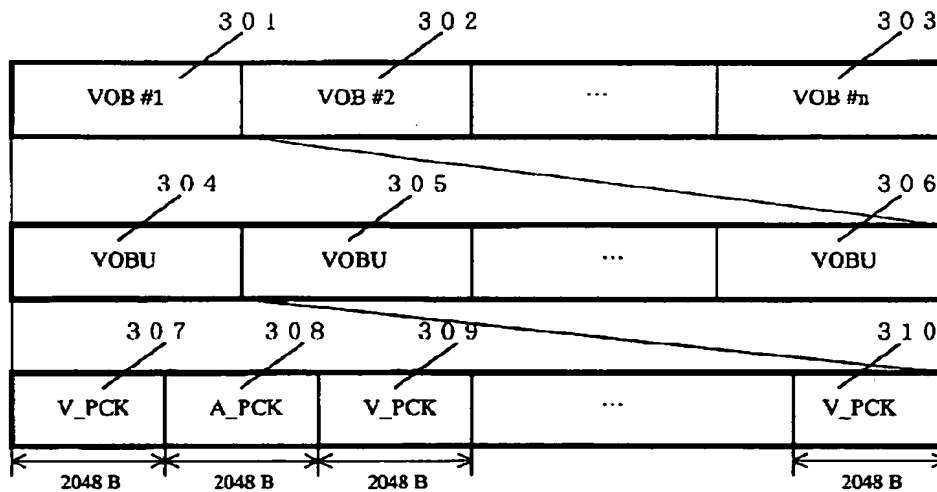
【図1】

図1



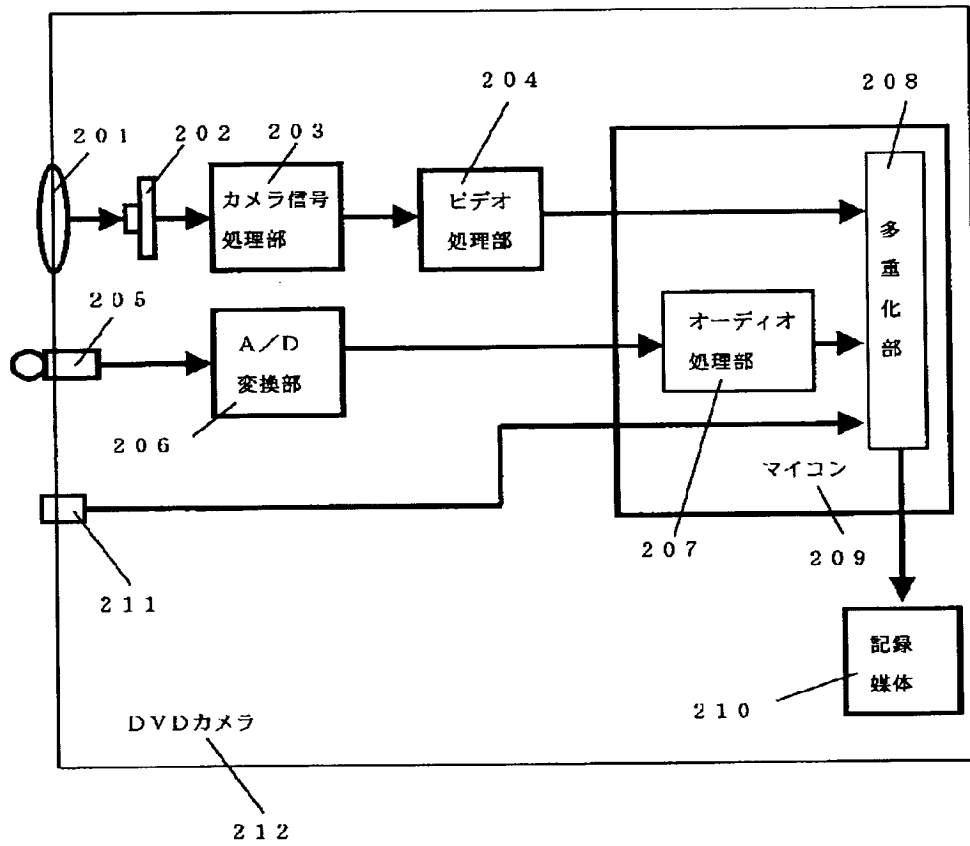
【図3】

図3



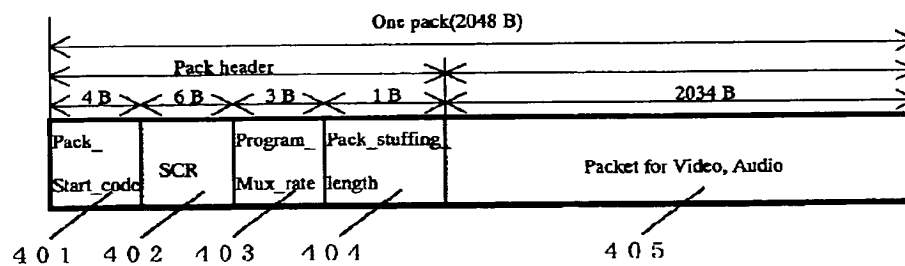
【図 2】

図 2



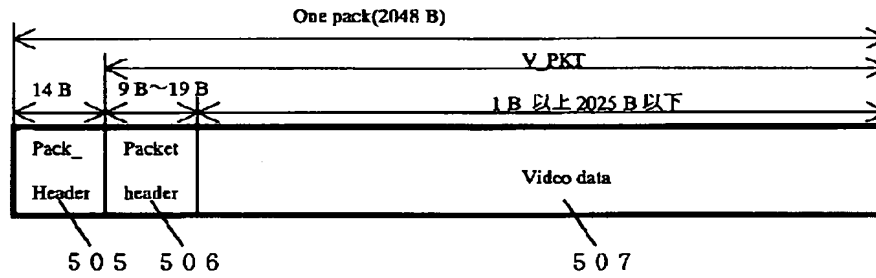
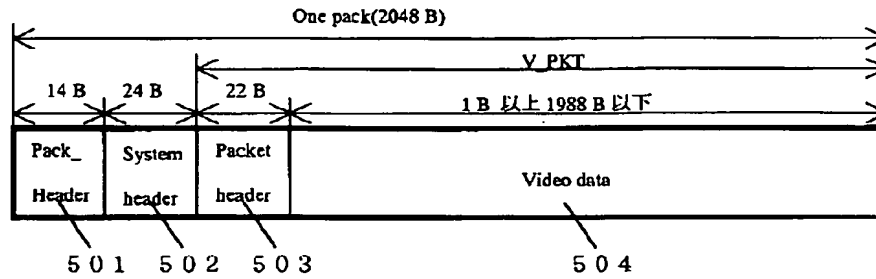
【図 4】

図 4



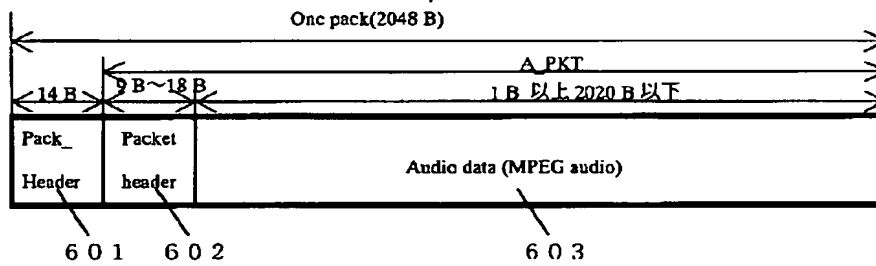
【図 5】

図 5



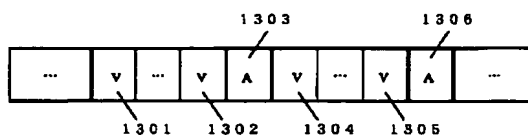
【図 6】

図 6



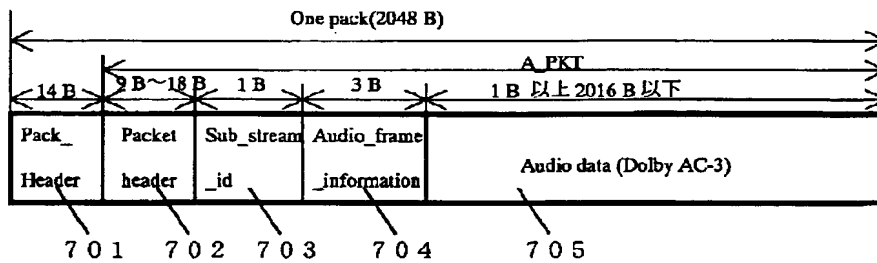
【図 13】

図 13



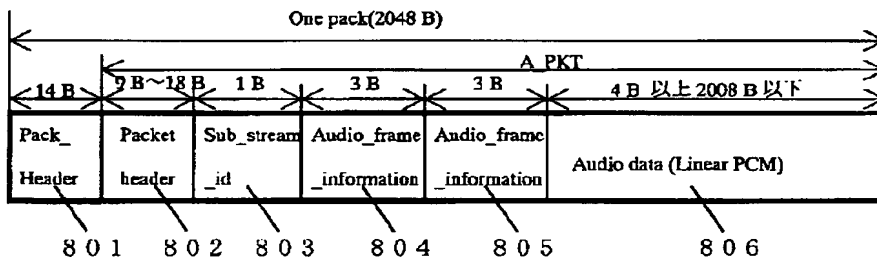
【図 7】

図 7



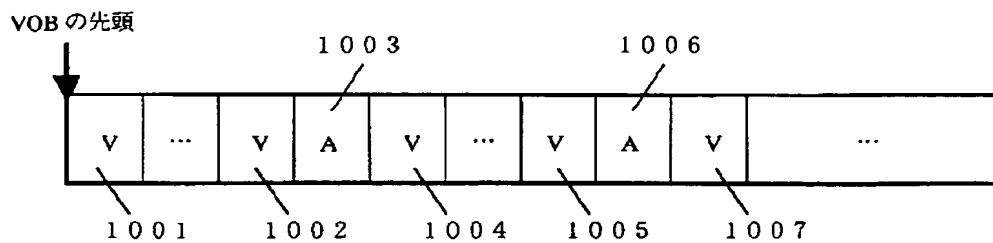
【図 8】

図 8



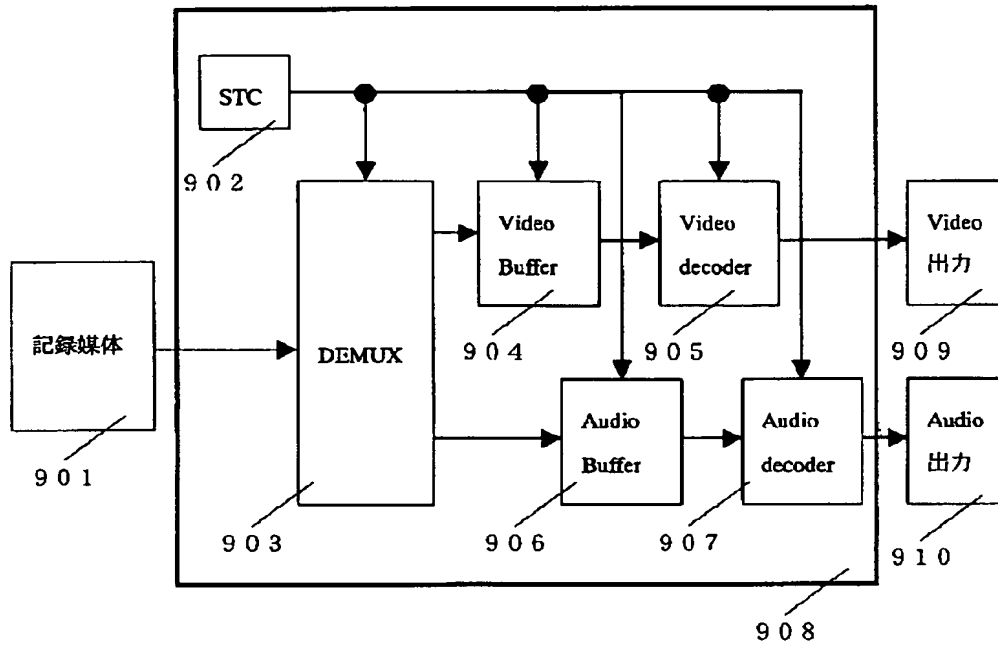
【図 10】

図 10



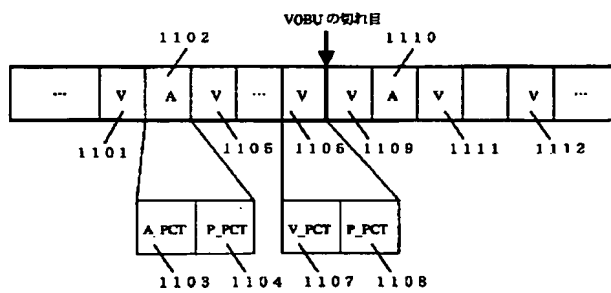
【図9】

図9



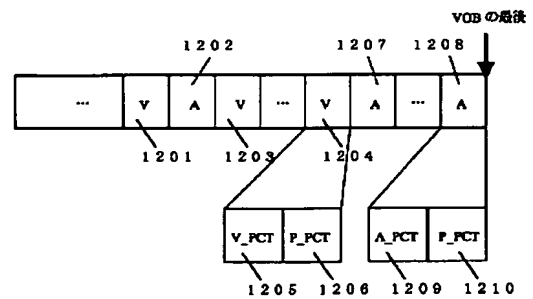
【図11】

図11



【図12】

図12



【図14】

図14

・ Video

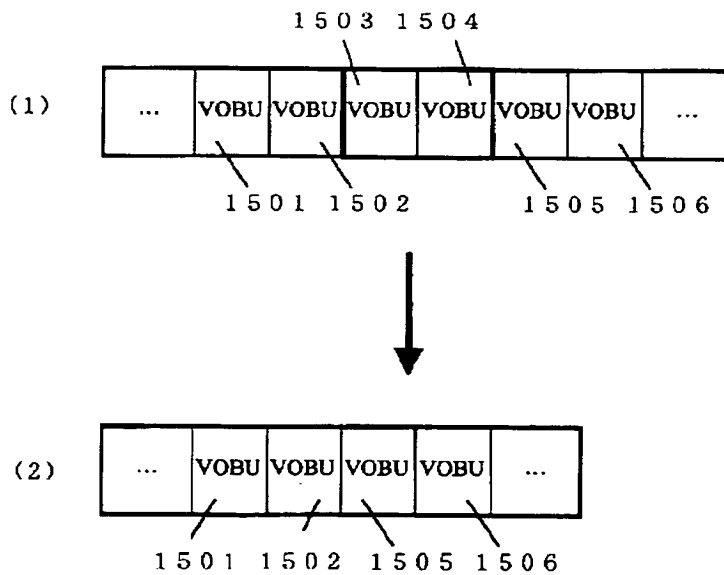
パターン	pack_header	system_header	packet_header	stuffing	padding_packet	total	Video Data
VOBU の先頭	14	24	22	0	0	60	1988
PTS,DTS あり	14	0	19	3	0	36	2012
PTS のみ	14	0	14	0	0	28	2020
PTS,DTS なし	14	0	9	1	0	24	2024
VOBU の最後	14	0	9~19	0~7	0 or more	24 or more	2025 or less

・ Audio

パターン	pack_header	system_header	packet_header	stuffing	padding_packet	total	Audio Data
VOBU の先頭	14	0	17	1	0	32	2016
PTS のみ	14	0	14	0	0	28	2020

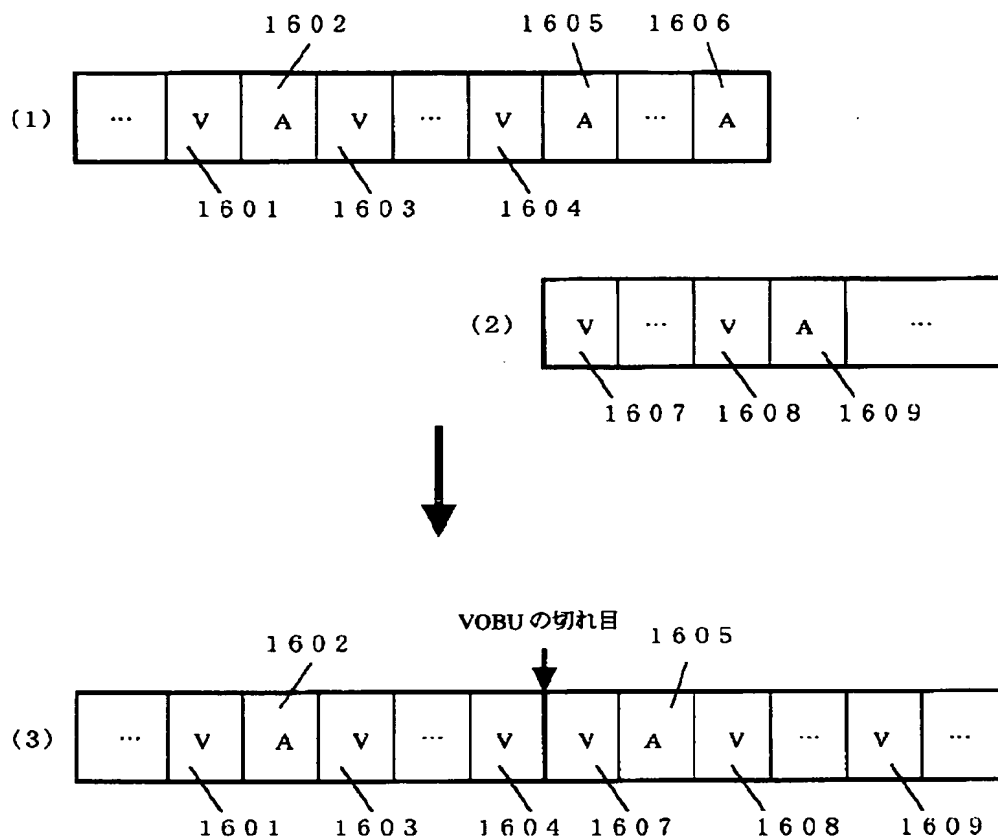
【図15】

図15



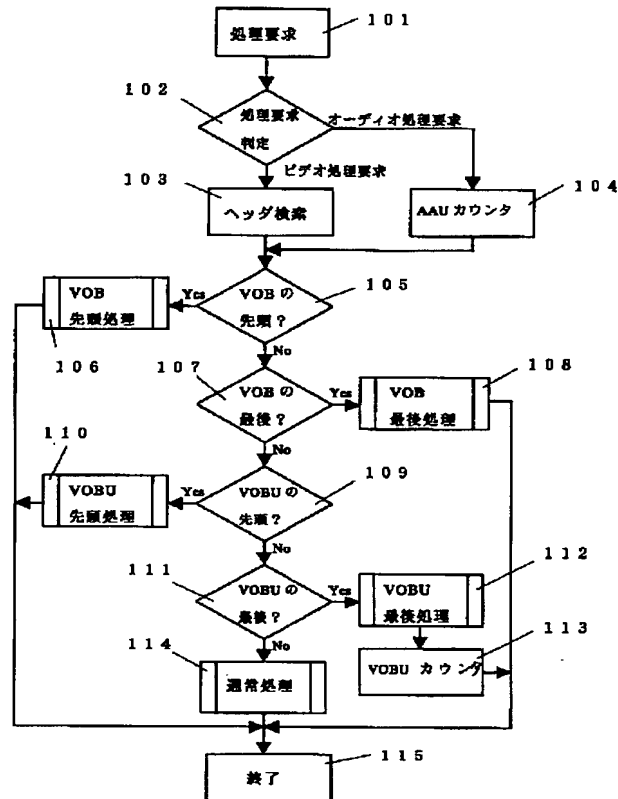
【図16】

図16



【図17】

図17



フロントページの続き

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
G 1 1 B 27/034		H 0 4 N 5/781	5 1 0 H
H 0 4 N 5/765			5 1 0 Z
5/781		5/91	L
5/91			N
5/92			C
7/24		5/92	H
		7/13	Z
		G 1 1 B 27/02	K

(72)発明者 垂水 浩幸
 神奈川県横浜市戸塚区吉田町292番地 株
 式会社日立画像情報システム内

F ターム(参考) 5C053 FA25 FA30 GB06 GB11 GB38
JA07 KA25 LA01 LA20
5C059 KK00 KK21 LA01 MA00 MA23
RB02 RC04 RC32 RC40 SS13
SS30
5D044 AB05 AB07 BC02 CC04 DE03
DE12 DE14 DE25 EF03 FG10
FG21 GK03 GK08 HL14
5D045 DA20
5D110 AA14 AA27 AA29 BB20 CA05
CA06 CA42 CK26

THIS PAGE BLANK (USPTO)